Class06

Jie

2022-10-14

Table of contents


```
df_grade <- read.csv("~/Desktop/BGGN 213/week3/student_homework.csv", row.names = 1)
df_grade
```

```
           hw1 hw2 hw3 hw4 hw5
student-1  100  73 100  88  79
student-2   85  64  78  89  78
student-3   83  69  77 100  77
student-4   88  NA  73 100  76
student-5   88 100  75  86  79
student-6   89  78 100  89  77
student-7   89 100  74  87 100
student-8   89 100  76  86 100
student-9   86 100  77  88  77
student-10  89  72  79  NA  76
student-11  82  66  78  84 100
student-12 100  70  75  92 100
student-13  89 100  76 100  80
student-14  85 100  77  89  76
student-15  85  65  76  89  NA
student-16  92 100  74  89  77
```

```
student-17  88  63 100  86  78
student-18  91  NA 100  87 100
student-19  91  68  75  86  79
student-20  91  68  76  88  76
```

```
df_grade[is.na((df_grade))] <- 0
df_grade
```

```
           hw1 hw2 hw3 hw4 hw5
student-1  100  73 100  88  79
student-2   85  64  78  89  78
student-3   83  69  77 100  77
student-4   88   0  73 100  76
student-5   88 100  75  86  79
student-6   89  78 100  89  77
student-7   89 100  74  87 100
student-8   89 100  76  86 100
student-9   86 100  77  88  77
student-10  89  72  79   0  76
student-11  82  66  78  84 100
student-12 100  70  75  92 100
student-13  89 100  76 100  80
student-14  85 100  77  89  76
student-15  85  65  76  89   0
student-16  92 100  74  89  77
student-17  88  63 100  86  78
student-18  91   0 100  87 100
student-19  91  68  75  86  79
student-20  91  68  76  88  76
```

Barry's demo

```
#grade <- function(x){
#x [is.na(x)] <- 0
#mean( x[-whcih.min(x)] )
```

```
#gradebook <- read.csv("~/Desktop/BGGN 213/week3/student_homework.csv",row.names = 1)
#head(gradebook)
```

'apply()' function apply(X, MARGIN, FUN, ..., simplify = TRUE)

```
#results <- apply(gradebook,1,grade)
#results
```

```
#which.max(results)
```

Question 1

```
df_grade["student-1",]
```

```
          hw1 hw2 hw3 hw4 hw5
student-1 100  73 100  88  79
```

```
s1 <- df_grade["student-1",]
s1
```

```
          hw1 hw2 hw3 hw4 hw5
student-1 100  73 100  88  79
```

```
df_grade[1,]
```

```
          hw1 hw2 hw3 hw4 hw5
student-1 100  73 100  88  79
```

```
vec = c()
grade <- function(df_grade) {
  for (i in 1:nrow(df_grade)){
    s <- df_grade[i,]
    average <- (rowSums(s) - min(s))/4

  vec = c(vec,average)
  }
    return(vec)
}
```

```
average_grade <- grade(df_grade)
average_grade
```

```
 student-1  student-2  student-3  student-4  student-5  student-6  student-7
     91.75      82.50      84.25      84.25      88.25      89.00      94.00
 student-8  student-9 student-10 student-11 student-12 student-13 student-14
     93.75      87.75      79.00      86.00      91.75      92.25      87.75
student-15 student-16 student-17 student-18 student-19 student-20
     78.75      89.50      88.00      94.50      82.75      82.75
```

```
#class(average_grade)
```

Question 2

```
which.max(average_grade)
```

```
student-18
        18
```

Question 3

```
mean_hw <- c(mean(df_grade$hw1),mean(df_grade$hw2),mean(df_grade$hw3),mean(df_grade$hw4),m
mean_hw
```

```
[1] 89.00 72.80 80.80 85.15 79.25
```

```
which.min(mean_hw)
```

```
[1] 2
```

Homework 2 is the hardest

Question 3 - sum

```
sum_hw <- c(sum(df_grade$hw1),sum(df_grade$hw2),sum(df_grade$hw3),sum(df_grade$hw4),sum(df
sum_hw
```

```
[1] 1780 1456 1616 1703 1585
```

```
which.min(sum_hw)
```

```
[1] 2
```

```
hw <- apply(df_grade,2,sum,na.rm=TRUE)
hw
```

```
 hw1  hw2  hw3  hw4  hw5
1780 1456 1616 1703 1585
```

Question 4

```
cor_5 <- cor(df_grade$hw5,average_grade)
cor_5
```

```
[1] 0.6325982
```

```
apply(df_grade,2,cor,y=average_grade)
```

```
      hw1       hw2       hw3       hw4       hw5
0.4250204 0.1767780 0.3042561 0.3810884 0.6325982
```

Homework 5 has the highest correlation